Working with Students to Reproduce COVID-19 Research to Establish the Credibility of Findings and Accelerate Policymaker Adoption

Dr. Peter Kedron | Arizona State University

Dr. Joseph Holler | Middlebury College

Project Objectives

1. Facilitate the adoption of reproducible research practices by GSI Fellows

2. Conduct reproductions and replications of GSI Fellows' research

3. Create educational materials to accelerate the adoption of reproducible research practices in the geographical sciences

Reproducibility

Obtaining consistent results using the same input data, computational steps, methods and code, and conditions of analysis

(National Academy of Sciences 2019)

Claerbout 1992; Donoho et al. 2009; Peng et al. 2006; Peng 2011

Replicability

Obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data
(National Academy of Sciences 2019)

Claerbout 1992; Donoho et al. 2009; Peng et al. 2006; Peng 2011

Table 2: Grouping of terminologies, as in Table 1, but by discipline.

A	B1	B2
political science	signal processing	microbiology, immunology (FASEB)
economics	scientific computing	computer science (ACM)
	econometry	
	epidemiology	
	clinical studies	
	internal medicine	
	physiology (neuro)	
	computational biology	
	biomedical research	
	statistics	

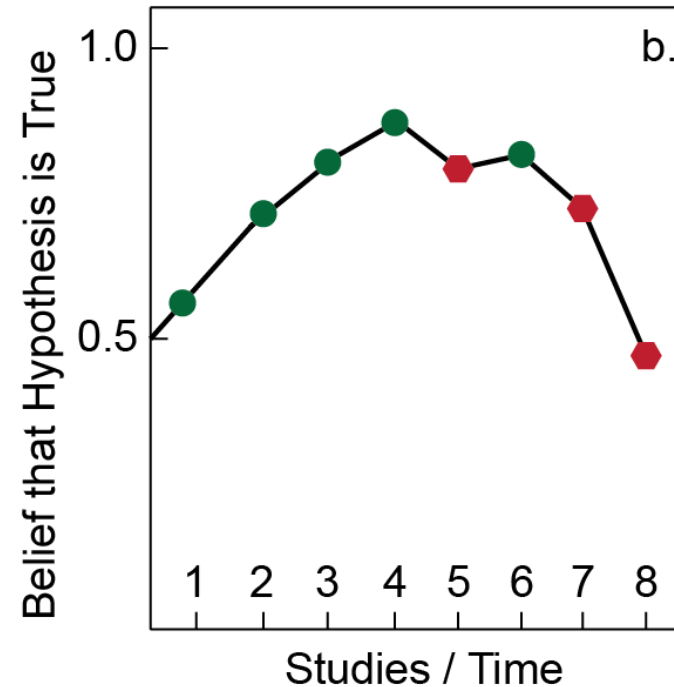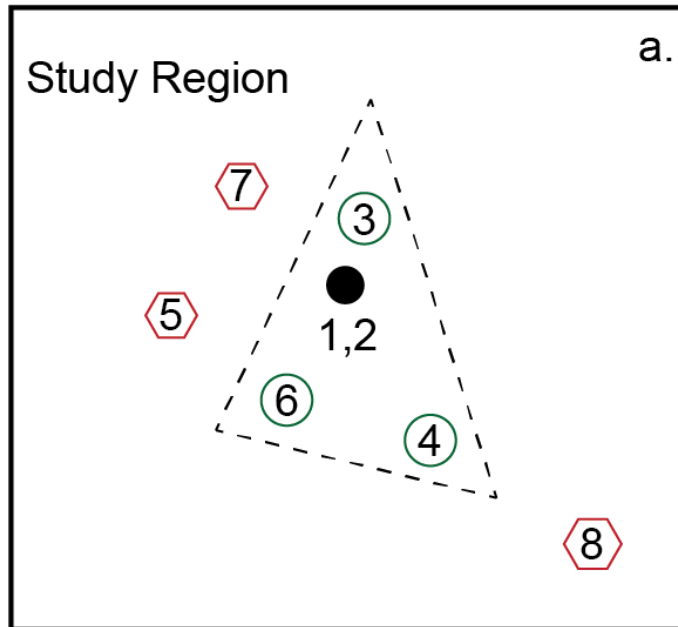Self-Correcting Mechanisms

R&R as a Means to an End

'Science builds explanatory structures, tells stories which are scrupulously tested to see if they are stories about real life.'
(Medawar 1967)

'Science is about finding the most reliable way of thinking at the present level of knowledge.'
(Rovelli 2014)

R&R as a Means to an End

The Need for R&R in COVID Research

- Knowledge of COVID-19 is *evolving rapidly*

- Public health depends upon *reliable* COVID-19 research

- Open research accelerates discovery
    - 9,698 articles on medRxiv and bioRxiv preprint servers (10/19/2020)
    - Tracking cases, deaths, government interventions, mobility
    - Accelerated peer review and open access publishing

- Retractions & perils due to *irreproducible data*
    - Hydroxychloroquine, Surgisphere, *The Lancet* & *New England Journal of Medicine*
    - location metadata missing for 68% of >12,000 genomic experiments (Schriml et al 2020)

- *Reproductions* establish *reliability*
    - weigh alternative explanations, meta-analysis cases
    - make and implement better policies

Many Necessary Conditions
Few Sufficient Conditions

- Many things are likely necessary for a piece of research to be reproducible

  - Available data and code, clear analytical framework, proper reporting

- Few (if any) things are sufficient for a piece of research to be reproducible

- Necessary and sufficient depend on context

- All this is even more complicated for replication

Conducting R&R:
Computational Infrastructure

- CyberGISX!

- Virtual computational environments
    - Reproducible hardware & software versions

- Data repository
    - Data & materials: open, public, documented metadata

- Jupyter notebooks
    - Code & methods: raw data → results with research code and narrative
    - Laboratory/field notebook for computational science

Conducting R&R:
Social & Institutional Infrastructure

- Pre-analysis planning:
  - theoretical constructs & hypotheses → data & methods

- Open publishing:
  - editorial standards, software standards
  - from pre-analysis through final paper, reviews
  - supplemental materials: data, software, code, apps
    - with clear citations / DOI's

- Education:
  - conducting & publishing reproducible research
  - conduct & publish reproduction & replication studies
  - reproductions & replications as pedagogy

Project Objectives

1. Facilitate the adoption of reproducible research practices by GSI Fellows

   a. Build a template Jupyter notebook

   b. Get to know your projects and plans

2. Conduct reproductions and replications of fellows' research

   a. Support interested fellows to prepare pre-analysis plans

   b. Publish white paper(s) of reproductions/replications

3. Create educational materials to accelerate the adoption of reproducible research practices in the geographical sciences

   a. Workshop series on reproducible research practices

   b. Develop teaching modules based on fellows' research